AAPM MEDICAL PHYSICS PRACTICE GUIDELINE # 5: Commissioning and QA of Treatment Planning Dose Calculations: Megavoltage Photon and Electron Beams

- Jennifer Simboketz (Chair), University of Wisconsin-Madison
- Indra Dae, Indiana University School of Medicine
- Vladimir Fougelman, Moffitt Cancer Center
- Benedick Fraass, Cedars-Sinai Medical Center
- Mark Geurts, University of Wisconsin-Madison
- Stephen Kry, MD Anderson
- Ingrid Marshall, Medical University of South Carolina
- Dimitra Mihaliotis, Charleston Radiation Therapy Center
- Zoubir Ouhib, Lynne Regional Cancer Center
- Timothy Ritter, University of Michigan
- Michael Snyder, Wayne State University
- Lynne Fairobent, AAPM Staff

AAPM Spring Clinical Meeting, Denver March 2014

What to do/check?

This report only cover dose calculation, the term "commissioning" includes beam data acquisition, modeling, and validation.

Introduction

Outline: follow outline of MPPG (plus rationale & some implementation experiences)

1. Introduction
   a. Goals
   b. Tolerances and evaluation criteria
   c. Scope/exclusions
2. Staff qualifications
3. Data acquisition
4. Model within TPS software
5. Photon beams: basic dose algorithm validation
   - MatLab code for TPS gamma analysis
   - Trilogy absolute dose verification, large field/off axis MLC tests
   - TomoTherapy: "tomographs"
6. Photon beams: heterogeneity correction validation
   - Clinac: CBCT phantom
7. Photon beams: IMRT/VMAT dose validation
   - TomoTherapy - TG 140 tests and clinical case
8. Electron beams
9. Routine QA (downloadable datasets)

Goals

While the implementation of robust and comprehensive QA programs recommended in other AAPM reports is strongly encouraged, the overall objective of this MPPG is to provide an overview of the minimum requirements for TPS dose algorithm commissioning and QA in a clinical setting. Specific goals for this report are to:

- Clearly identify and reference applicable portions of existing AAPM reports and peer-reviewed articles for established commissioning components.
- Provide updated guidelines on technologies that have emerged since the publication of previous reports.
- Provide guidance on validation tests for dose accuracy and constancy (select downloadable datasets/contours & beam parameters are provided for optional use).
- Provide guidance on typical achievable tolerances and evaluation criteria for clinical implementations.
- Provide a checklist for commissioning processes and associated documentation.

Scope/exclusions

- Title: Commissioning and QA of Treatment Planning Dose Calculations: Megavoltage Photon and Electron Beams
- The scope of this report is limited to the commissioning and QA of the beam modeling and calculation portion of a TPS where:
  - External photon and electron treatment beams are delivered at typical SSDs using a Linac mounted radiation source including conventional and small fields used in IMRT, VMAT, helical tomotherapy delivery, and SRS/SBRT (all up for discussion).
  - Modern dose algorithms are utilized including corrections for tissue heterogeneity.
  - The Multi-Leaf Collimator (MLC) is used as the primary method of shaping the beam aperture for treatments. (Individually fabricated IMRT modifiers, cones... still up for discussion)
- Excludes: (not an exhaustive list, and not all written in document)
  - Non-dosimetric components of system, e.g.: DVH, leaf sequences, contours, image registration.
  - Brachytherapy.
  - Proton therapy.
  - Non-commercial planning systems.
  - Radiation delivered by robots.

Tolerances & Evaluation Criteria (2 "tier approach")

- Wanted to state minimum acceptable tolerance for TPS "basic" dose calculation:
  - "The tolerances for the basic photon tests are widely accepted as minimum criteria for static photon beams under conditions of charged particle equilibrium."
- Wanted to push the limit on some evaluation criteria to find limitations of dose calculations:
  - "Given that there is not widely accepted minimum tolerance for the other verification tests in this MPPG, (including those for VMAT/IMRT), those evaluation criteria must not be interpreted as mandatory or regulatory tolerances. Rather, they are values defined as points for further investigation, possible improvement, and resolution."
- Did not want to state or use any minimum tolerance values not widely accepted/published:
  - "All the tolerances and criteria in this report are based on a combination of published guidelines, the dosimetric audits performed by the Radiological Physics Center, and the experience of authors. Users are encouraged to not only meet these tolerances, but also strive to achieve dosimetric agreement comparable to that reported in the literature for their particular algorithm."
Outline: follow outline of MPPG (plus rationale & some implementation experiences)

1. Introduction
   a. Goals
   b. Tolerances and evaluation criteria
   c. Scope/exclusions
2. Staff qualifications
3. Data acquisition
4. Model within TPS software
5. Photon beams: basic dose algorithm validation
   - Matlab code for 1D gamma analysis
   - Trilogy: absolute dose verification, large field/offset axis MLC tests
   - TomoTherapy: "tomophants"
6. Photon beams: heterogeneity correction validation
   - Clinac: CIRMS phantom
7. Photon beams: IMRT/VMAT dose validation
   - TomoTherapy – TG 119 tests and clinical case
8. Electron beams
9. Routine QA (downloadable datasets)

Data Acquisition Question

What data do you use when commissioning a new dose algorithm?

1. Collect data according to vendors guidelines
2. Collect some of the vendor recommended data but not all
3. Collect all required data and more
4. Use golden beam data
5. Hey, I thought this wasn’t a SAM session.

Staff, Data, Model…

- Staff qualifications – QMP, defer to supervision MPPG
- Data acquisition – defer to TPS manuals for all required data (water tank, and in air for MC) & refer to TG 106. An equipment list/summary on small field/MLC data acquisition is included:
  - PDD and QF with a small volume detector down to at least 2x2 cm²
  - MLC intra and inter-leaf transmission and leaf gap:
    - Large chamber if an average intra- and inter-leaf value is specified.
    - Separate measurements, use small chamber under the leaf and film for inter-leaf leakage measurements
  - Measure leaf-end penumbra with a small detector (such as a diode or micro-chamber) to avoid volume-averaging effects
  - Leaf timing for binary MLC systems should be verified using film or exit detector measurements
- Model – refer to manual, iterate as needed using results from validation testing

Validation Question

What type of dose algorithm validation do you do as part of the commissioning process?

1. None
2. Routine patient specific DQA serves as validation
3. In-house test suite (chamber, array, films etc…)
4. Peer review audit (colleague or RPC)
5. Combination of 3 and 4

Validation Measurements

- Water tank, ion chambers & diodes
- Custom phantom
- IMRT DQA Device (i.e. Delta4)

- Report was written such that user has freedom to use any suitable/available combination of phantoms and detectors.
- Combination of in-house and external audits
- It is recommended to take data at time of commissioning
- This diagram shows a common set of tools (and what we are using at UW.)
5. Basic Validation: Photon beams

Section 5 (Photons in homogeneous media) has 2 sets of tests:
- 5.1-5.3: "sanity check" of commission data → physics module → planning module and TG 51 calibration value
- 5.4-5.9: test fields that were not used in commissioning. Compare measured and calculated dose distribution.

Tests should be run for each unique configured beam (energy and wedge).

Implementation: 5. Dose in test plan vs. TPS calibration (0.5% tolerance)

<table>
<thead>
<tr>
<th>Test</th>
<th>Description</th>
<th>Tolerance</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.1</td>
<td>Dose distribution in planning module vs. physics module</td>
<td>--</td>
</tr>
<tr>
<td>5.2</td>
<td>Dose in test plan vs. clinical calibration condition</td>
<td>0.5%</td>
</tr>
<tr>
<td>5.3</td>
<td>Dose distribution calculated in planning system vs. commissioning data</td>
<td>2%</td>
</tr>
</tbody>
</table>

No additional measurements beyond commissioning data needed for these tests.

Accuracy question

How accurate is your worst off axis relative dose calculation?
1. 1%
2. 2%
3. 3%
4. 4%
5. 5%

Accuracy question 2

How accurate is your worst off axis relative dose calculation with a wedge in place?
1. 1%
2. 2%
3. 3%
4. 4%
5. 5%
Section 5: Basic photon tolerances

Table 5. Basic TPS photon beam evaluation methods and tolerances

<table>
<thead>
<tr>
<th>Region</th>
<th>Evaluation Method</th>
<th>Tolerance* (consistent with RPC)</th>
</tr>
</thead>
<tbody>
<tr>
<td>High dose</td>
<td>Relative dose with one parameter change from reference conditions</td>
<td>2%</td>
</tr>
<tr>
<td></td>
<td>Relative dose with multiple parameter changes.**</td>
<td>3%</td>
</tr>
<tr>
<td>Penumbra</td>
<td>Distance to agreement</td>
<td>3 mm</td>
</tr>
<tr>
<td>Low dose tail</td>
<td>Up to 3 cm from field edge</td>
<td>3% of maximum field dose</td>
</tr>
</tbody>
</table>

* Tolerances are relative to local dose unless otherwise noted.
**e.g. off axis with physical wedge.

Implementation: 5.5 Large MLC shaped field with extensive blocking ($\gamma$ analysis)

- Example of a test pattern – that tests many things at once: Off axis PDD (+), 3 cross profiles (2 cm, 10 cm, 20 cm) and 1 in line profile (10 cm) for open and wedge fields.

60° wedge, "toe in"

1D Gamma analysis– open source MatLab code

- Save scan data in Excel and output dicom dose files from TPS (note dose grid origin and resolution).
- Script/detailed users manual will be available on the UW Open Source Medical Devices website and code revision history at github:
  - https://github.com/bredfeldt/MPPG

Thanks to MatLab Master Jeremy Bredfieldt!

Gamma Calculation Test Case

Min. $\gamma$ will occur with a dose error is 0.015 v($\sqrt{2}$) and position error is 1.5 v($\sqrt{2}$)

Gamma Calculation Test Results

Validate gamma calculation with 3%/3mm threshold:
- Create simulated dose profiles A and B
  - $A = \text{dose ramp with slope} = 0.03 \text{ Gy/3mm}$
  - $B = A + 0.03\sqrt{2}$
- Input A and B into gamma calculation
- Verify that gamma = 1 at all positions
Results from off axis PDD for open 10MV field
2%/0.001mm

Problem in buildup region. Adjust model of the electron contamination?

Results from off axis PDD for 15 wedge 10MV field
5%/0.001mm

Problem in buildup region. Adjust model of the electron contamination?

Results from d=10 cm inline profile for 30° wedged 10MV field, γ = 3%/3mm

1. Problem in penumbra (T&G) region. Adjust leaf intra or inter leaf leakage model?
2. Problem with jaw closing to MLC defined edge?

Results from d=10 cm inline profile for 30° wedged 10MV field, γ = 5%/3mm

2 parameters change (off-axis, and wedge)

SCAN DICTION

1. Problem in penumbra (T&G) region. Adjust leaf intra or inter leaf leakage model?
2. Problem with jaw closing to MLC defined edge?

Results from d=25 cm crosline profile for 60° wedged 10MV field, γ = 3%/3mm

Implementation: Basic test on TomoTherapy

- Forward planned fields are not easily generated in tomo
- "TomoPhants": set of standard plans with different jaw sizes (fixed and dynamic) run on "cheese phantom" with ion chambers for inline profiles (and Delta4 for volumetric DQA.) They are a good alternative to implementation of section 5 tests
- Calculated dose profiles are extracted by Accuray and measured data is analyzed with excel sheets

Implementation: Basic test on TomoTherapy

- Forward planned fields are not easily generated in tomo
- "TomoPhants": set of standard plans with different jaw sizes (fixed and dynamic) run on "cheese phantom" with ion chambers for inline profiles (and Delta4 for volumetric DQA.) They are a good alternative to implementation of section 5 tests
- Calculated dose profiles are extracted by Accuray and measured data is analyzed with excel sheets
Implementation: TomoPhant results

- Same plan with helical and tomoelectric, for each field size. Results:
  - TomoDirect measured hotter than planned (compared to helical plans)
  - 5 cm FW plans were always hotter than planned (compared to other FW)
  - A1SL ion chamber, error bars are 3%
- What can be done? Adjust JFOF (basically a collimator scatter output factor (Sc) table)

![Graph showing differences in temperature]

Max error = 1.2%

Heterogeneity questions

Which algorithm is not acceptable for dose calculation for lung?
1. Pencil beam
2. Monte Carlo
3. Convolution superposition
4. Discreet ordinance (grid based Boltzman solver)
5. All are acceptable

Section 6: Heterogeneity

<table>
<thead>
<tr>
<th>Test</th>
<th>Objective</th>
<th>Description (example)</th>
<th>Detector</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Verify small field PDD</td>
<td>≥3 cm MLC shaped fields, with PDD acquired at a clinically relevant SSD</td>
<td>Density plane scanner</td>
</tr>
<tr>
<td>1.2</td>
<td>Verify output for small MLC defined fields</td>
<td>Use small square and rectangular MLC-defined shapes, measuring output at a clinically relevant depth for each*</td>
<td>Density plane scanner, ion chamber or measurement chamber</td>
</tr>
<tr>
<td>1.3</td>
<td>TG-141 ion</td>
<td>Plan, measure, and perform planning and QA tests in the TG-141 report for both the fixed and dynamic MLC shapes</td>
<td></td>
</tr>
<tr>
<td>1.4</td>
<td>Clinical tests</td>
<td>Choose at least one clinical case, plan, measure, and perform an in-depth analysis of the results.</td>
<td>Ion chamber, film, and/or array</td>
</tr>
</tbody>
</table>

Section 7: IMRT/VMAT Verification

<table>
<thead>
<tr>
<th>Test</th>
<th>Objective</th>
<th>Description (example)</th>
<th>Detector</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Verify small field PDD</td>
<td>≥3 cm MLC shaped fields, with PDD acquired at a clinically relevant SSD</td>
<td>Density plane scanner</td>
</tr>
<tr>
<td>1.2</td>
<td>Verify output for small MLC defined fields</td>
<td>Use small square and rectangular MLC-defined shapes, measuring output at a clinically relevant depth for each*</td>
<td>Density plane scanner, ion chamber or measurement chamber</td>
</tr>
<tr>
<td>1.3</td>
<td>TG-141 ion</td>
<td>Plan, measure, and perform planning and QA tests in the TG-141 report for both the fixed and dynamic MLC shapes</td>
<td></td>
</tr>
<tr>
<td>1.4</td>
<td>Clinical tests</td>
<td>Choose at least one clinical case, plan, measure, and perform an in-depth analysis of the results.</td>
<td>Ion chamber, film, and/or array</td>
</tr>
</tbody>
</table>

IMRT DQA Question 1

What gamma criteria do you use for patient specific delivery QA (DQA)?
1. 1%/1mm
2. 2%/2mm
3. 3%/3mm
4. 4%/4mm
5. I don’t do patient specific DQA and/or I don’t use gamma criteria for DQA analysis.
What do you do when a case ‘fails’ that criteria?
1. Increase tolerance by 1%/1mm
2. Re-measure
3. Re-plan
4. Pick tolerance so >95% pass and report tolerance values
5. My plans never fail

---

**IMRT/VMAT Validation Tests (section 7)**

**TG 119 C-shaped plan on tomo with Delta4**

- Delta4 2%/2mm (global) gamma analysis
- Use only detectors with >20% signal
- Excellent results, 100% pass

---

**Downloadable data sets with plan instruction**

---

**Section 8: Electron Beam Verification**

<table>
<thead>
<tr>
<th>Test</th>
<th>Objective</th>
<th>Description</th>
<th>Tolerance</th>
</tr>
</thead>
<tbody>
<tr>
<td>8.1</td>
<td>Basic model verification with shaped fields</td>
<td>Custom cones at standard and extended SSDs</td>
<td>3%/3 mm</td>
</tr>
<tr>
<td>8.2</td>
<td>Surface irregularities-obliquity</td>
<td>Oblique incidence using reference cone and nominal clinical SSD</td>
<td>5%</td>
</tr>
<tr>
<td>8.3</td>
<td>Inhomogeneity test</td>
<td>Reference cone and nominal clinical SSD</td>
<td>7%</td>
</tr>
</tbody>
</table>
Section 9 QA

- Annually or after major TPS upgrades
- Reference plans should be selected at the time of commissioning and then re-calculated for routine QA comparison.
- For photons, representative plans for each configured beam should be chosen from Table 4 for static and wedge beams and Table 7 for IMRT/VMAT. Optionally, an additional thorax dataset with contours and suggested static beam parameters can be downloaded and used for some of these tests. ([http://www.aapm.org/pubs/tg244/](http://www.aapm.org/pubs/tg244/)). A 10x10 cm$^2$ field and a small field (e.g. 5x5 cm$^2$) can be prescribed to the isocenter located in the center of the PTV. Wedged fields and dynamic arc plans can also be calculated on the thorax data set.
- For electrons, plans should be calculated for each energy using a heterogeneous dataset with reasonable surface curvature. The sample thorax dataset is also suitable for this test. Recommended plans also include extended distance and bolus verification.
- The routine QA re-calculation should agree with the reference dose calculation to within 1%/1mm. A complete re-commissioning (including validation) may be required if more significant deviations are observed.

Next steps….

- Respond to public comment reviewer comments
- Submit to JACMP – await final review
- Continue implementation of MPPG on Varian, TomoTherapy and Elekta (AAPM annual meeting abstract)
  - Fine tune gamma analysis in Matlab code, analyze remaining Trilogy and Infinity data
  - Take heterogeneous and electron data
  - Create test suite for each machine type (Pinnacle/Eclipse plans, R&V entry and scan Q’s)
- Make gamma analysis code easily available (and easier data input)

Checklist to guide commissioning report

<table>
<thead>
<tr>
<th>Section</th>
<th>TG244 Note</th>
<th>Consequence Report Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>QF-5 endorses algorithms and has received peer review.</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Manufacturer’s guidance for data acquisition was corrected and clarified.</td>
<td></td>
</tr>
<tr>
<td>5.0</td>
<td>Reference 2D calibration data included.</td>
<td></td>
</tr>
<tr>
<td>5.0</td>
<td>Reference 3D calibration data included.</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Beam modeling process completed according to vendor’s recommendations.</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Beam modeling process completed according to vendor’s recommendations.</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>For each beam type, perform calculation tests 5.5.3.2.5 and 5.5.3.2.6 for non-physical wedges (e.g., wedge factor).</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Heterogeneity corrections validated for photon beams.</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>IMRT and VMAT calculations accomplished for each configured beam according to Table 9.</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>IMRT and VMAT calculations accomplished for each configured beam according to Table 9.</td>
<td></td>
</tr>
<tr>
<td>7.0</td>
<td>Underdose and overdose limitations of IMRT/VMAT modeling and dose calculations.</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Electron validation performed according to Table 9.</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Baseline (0,0) protocol for each machine type identified for each configured beam and routine QA established.</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>Four more obtained and any recommendations addressed.</td>
<td></td>
</tr>
</tbody>
</table>

Thanks to my collaborators, and to you for your attention!

- All MPPG#5 members!
- UW clinical physicists who helped with implementation
  - Adam Bayliss
  - John Bayouth
  - Ed Bender
  - Jessica Miller
- UW Medical Physics graduate students
  - Jeremy Bredtfeldt
  - Sam Simiele